

Adaptive multisensor fusion using the first two moments of the innovation under nonstationary degradation

Fusión multisensor adaptativa mediante los dos primeros momentos de la innovación bajo degradación no estacionaria

F. J. Burgos-Flórez  ; J. G. Popayán-Hernández  ; J. P. Hoyos-Sánchez 

DOI: <https://doi.org/10.22517/23447214.26509>

Scientific and technological research paper

Abstract— Adaptive Kalman filters modify the confidence assigned to measurements when sensor quality changes. Many methods monitor innovation energy, although a slowly varying bias can shift its mean without producing large instantaneous innovations. This work proposes a channel-wise adaptation that combines bounded covariance inflation, a test on the mean of the whitened innovation, and a minimum-dwell-time gate. A differential-drive ground vehicle with GNSS, odometry, gyroscope, and magnetometer was simulated. Six conditions, five severity levels, and forty paired repetitions per cell were studied. A nominal EKF, a mistuned EKF, covariance matching, Sage-Husa, robust Huber weighting, and the proposed method were compared. Position RMSE predicted tracking RMSE with a mean Spearman correlation of 0.994, but the association fell to 0.491 for recovery time. Second-moment methods responded to variable noise and outliers but did not correct drifting bias. The first-moment test enabled action against this shift, although exclusion exhibited a bimodal distribution and improved 28 of 40 repetitions. Minimum dwell time reduced nominal switching. At maximum severity, test-triggered compensation improved upon the nominal EKF in all 40 repetitions, with a median improvement of 67.0% and a nominal cost of 32.2%. No action simultaneously dominated nominal performance, median bias performance, and dispersion. Adaptation should distinguish variance and mean changes and should be evaluated using average, transient, and nominal-cost metrics.

Index Terms— adaptive Kalman filter; closed-loop control; multisensor fusion; sensor degradation; state estimation.

Resumen Los filtros de Kalman adaptativos modifican la confianza asignada a las mediciones cuando cambia la calidad de los sensores. Muchos métodos monitorean la energía de la innovación, aunque un sesgo de variación lenta puede desplazar su media sin producir innovaciones instantáneas grandes. Este trabajo propone una adaptación por canal que combina inflación de covarianza acotada, una prueba sobre la media de la innovación blanqueada y una compuerta con tiempo mínimo de permanencia.

This manuscript was submitted on September 02, 2026, accepted on September 16, 2026, and published on September 30, 2026. This work was supported by the Universidad Nacional de Colombia, Sede La Paz, through the project “Desarrollo y evaluación de estrategias de modelado, estimación y control para sistemas mecatrónicos y robóticos” (code 69296).

F. J. Burgos-Flórez is a Ph.D. in Mechanical Engineering and is with the Escuela de Pregrado, Dirección Académica, Vicerrectoría de Sede, Universidad Nacional de Colombia, Sede La Paz, La Paz, Cesar, Colombia, e-mail: fjburos@unal.edu.co.

Se simuló un vehículo terrestre de tracción diferencial con GNSS, odometría, giróscopo y magnetómetro. Se estudiaron seis condiciones, cinco niveles de severidad y cuarenta repeticiones pareadas por celda. Se compararon un EKF nominal, un EKF mal sintonizado, covarianza matching, Sage Husa, ponderación robusta de Huber y el método propuesto. El RMSE de posición predijo el RMSE de seguimiento con una correlación media de Spearman de 0.994, pero la asociación disminuyó a 0.491 para el tiempo de recuperación. Los métodos de segundo momento respondieron al ruido variable y a los valores atípicos, pero no corrigieron el sesgo con deriva. La prueba de primer momento permitió actuar ante este desplazamiento, aunque la exclusión presentó una distribución bimodal y mejoró 28 de 40 repeticiones. El tiempo mínimo de permanencia redujo la conmutación nominal. A máxima severidad, la compensación activada por la prueba mejoró al EKF nominal en las 40 repeticiones, con una mejora mediana de 67.0% y un costo nominal de 32.2%. Ninguna acción dominó simultáneamente el desempeño nominal, la mediana bajo sesgo y la dispersión. La adaptación debe distinguir los cambios de varianza y de media, y debe evaluarse mediante métricas promedio, transitorias y de costo nominal.

Palabras claves— control en lazo cerrado; degradación de sensores, estimación de estado; filtro de Kalman adaptativo; fusión multisensor.

I. INTRODUCTION

AUTONOMOUS systems estimate their state by integrating a process model with measurements from sensors that differ in rate, units, and reliability. This integration becomes more difficult when sensing quality changes during operation [1]. Within the Kalman framework, the process model propagates the state and its uncertainty, whereas the observation model incorporates each sensor measurement [2]. The balance between these two information sources is governed by the process and measurement noise covariances, which are often

J. G. Popayán-Hernández is a Ph.D. in Environmental Sciences and is with the Departamento de Ciencias Agrícolas, Facultad de Ciencias Agropecuarias, Universidad Nacional de Colombia, Sede Palmira, Palmira, Valle del Cauca, Colombia, e-mail: jgpopayanh@unal.edu.co.

J. P. Hoyos-Sánchez is a Ph.D. in Electronic Sciences and is with the Escuela de Pregrado, Dirección Académica, Vicerrectoría de Sede, Universidad Nacional de Colombia, Sede La Paz, La Paz, Cesar, Colombia, e-mail: jhoyoss@unal.edu.co.



treated as fixed despite changing operating conditions [3]. When these covariances are misspecified, the measurement gain and the reported uncertainty can become inconsistent with the true estimation error [4].

Adaptive estimation adjusts covariances during operation [1,2]. Consistency provides a basis for tuning the normalized estimation error squared (NEES) and normalized innovation squared (NIS) [1]. NIS drives adaptations in inertial arrays [2], batteries [3], and visual-inertial odometry [4]. The innovation can also be used for fault detection [5]. In mobile robotics, Student-t distributions [6], adaptation under degraded positioning [7], and eigenvalue-based indicators [8] have been used.

Fading factors respond to time-varying uncertainty [9]. Adaptive sigma-point filters address model error [10]. Consistency-based tuning includes Bayesian optimization [11] and underwater applications [12]. Measurement covariance has also been estimated through Bayesian inference [13] and robust autocovariance methods [14].

Innovation-based adaptation has been applied to inertial and satellite navigation [15]. It has also been used with redundant measurements [16], dual-threshold tests [17], and sensor and actuator faults [18]. Other robust variants address pedestrian heading [19], weak satellite signals [20], urban LiDAR [21], and outliers [22]. Additional applications include underwater navigation [23], federated fusion [24], redundancy outside navigation [25], abnormal-measurement tests [26], carrier tracking [27], and multiple-model banks [28]. This literature often treats dispersion, mean shift, and post-detection action separately.

Multivariate exponentially weighted means can accumulate small persistent shifts [29]. CUSUM schemes provide another classical formulation for sequential change detection [30]. Therefore, monitoring the first moment is not novel by itself. The gap addressed here is its channel-wise integration with covariance inflation, dwell time, exclusion or compensation, followed by a common evaluation in estimation and control.

The difference with respect to the comparators used in this study can be stated in terms of what each method monitors and how it responds. Covariance matching rescales the measurement covariance so that the observed innovation dispersion agrees with the predicted one [31]. Sage-Husa estimation recursively updates the measurement-noise level from the innovation sequence. Huber weighting bounds the influence of large residuals and therefore reduces the effect of outliers [32]. All three act on the second moment of the innovation and respond through a single mechanism, namely a change in the weight of the measurement. The proposed scheme differs in three respects. First, it monitors the mean of the whitened innovation, which none of these comparators tracks. Second, it separates the two signals: the energy statistic governs bounded inflation, whereas the mean statistic flags a persistent shift. Third, it couples detection with an explicit post-detection action, either channel exclusion or activation of a bias state, instead of only rescaling the covariance.

The first challenge is statistical. NIS summarizes the quadratic magnitude of the innovation. An increase in variance raises this statistic. A slowly varying bias can shift the mean while keeping instantaneous innovations moderate because the filter gradually follows the biased measurement. In that case, the second moment may be insufficient.

Throughout this work, first-moment information refers to the exponentially weighted mean of the whitened innovation, whereas second-moment information is represented by its normalized quadratic innovation energy through the NIS.

The second challenge concerns evaluation. Estimators are commonly compared using open-loop RMSE. The controller consumes the estimated state and may respond differently to smooth errors, bursts, or delays. Therefore, a lower RMSE does not necessarily guarantee a better transient response. This relationship should be measured rather than assumed.

The objective is to propose and evaluate an adaptive multisensor-fusion scheme that combines the first two moments of the innovation to respond differently to dispersion changes and persistent biases, and to quantify their effects on estimation and closed-loop control.

Three questions are addressed. Does estimation RMSE preserve the ranking of estimators when average and transient control performance are evaluated? Does statistical consistency predict average control performance better than RMSE? Does the combination of mean, energy, and dwell time enable differentiated responses to noise, outliers, and drifting bias?

The tested hypotheses were as follows. H1 stated that the ranking by estimation RMSE would not be preserved in closed loop. H2 stated that inconsistency measured through NEES would predict mean control error better than RMSE. H3 stated that a minimum dwell time would reduce spurious switching without eliminating the response to bias. Bias compensation was evaluated as a secondary analysis using a disjoint seed set.

The contributions are fourfold. First, a channel-wise algorithm is formalized that combines bounded inflation, a whitened exponentially weighted mean, and a dwell-time gate. Second, the study quantifies when RMSE predicts mean performance and when it loses information about the transient response. Third, an ablation identifies the contribution of the first-moment test under drifting bias. Fourth, the trade-off between exclusion, triggered compensation, and permanent compensation is quantified. The proposal integrates known components under a reproducible rule and a common evaluation with explicit nominal costs.

II. MATERIALS AND METHODS

A. Plant and Reference Trajectory

The simulated platform is a differential-drive ground vehicle moving in the plane. The true state was defined as

$$\mathbf{x} = [p_x, p_y, \psi, v, \omega, b_g]^T \quad (1)$$

where p_x and p_y are position, ψ is heading, v is longitudinal velocity, ω is angular velocity, and b_g is gyroscope bias. The input contains the linear- and angular-velocity commands v_c and ω_c . The continuous dynamics used the unicycle model with first-order lags in (2), where τ_v and τ_ω are the time constants of the linear and angular velocity, and w_b is the white noise that drives the gyroscope-bias random walk.

$$\begin{aligned} \dot{p}_x &= v \cos \psi, & \dot{p}_y &= v \sin \psi, & \dot{\psi} &= \omega \\ \dot{v} &= \frac{v_c - v}{\tau_v}, & \dot{\omega} &= \frac{\omega_c - \omega}{\tau_\omega}, & \dot{b}_g &= w_b \end{aligned} \quad (2)$$

Integration was performed at 100 Hz using the fourth-order Runge-Kutta method. The time constants were 0.35 s for linear velocity and 0.20 s for angular velocity. The diffusion coefficients were $0.08 \text{ ms}^{-\frac{3}{2}}$ for longitudinal velocity, $0.06 \text{ rad s}^{-\frac{3}{2}}$ for angular velocity, and $2 \times 10^{-4} \text{ rad s}^{-\frac{3}{2}}$ for the gyroscope-bias random walk. Commands were limited to 2.0 m/s and 2.5 rad/s. At the base step of 0.01 s, the simulator adds each coefficient multiplied by the square root of the time step to an independent standard normal variate, and the filter uses the corresponding discrete variance $\sigma_i^2 \Delta t$.

The reference was a Geronio lemniscate with 6 m amplitude and a 40 s period. Its longitudinal speed ranged from 0.623 to 1.333 m/s, and its maximum angular velocity was 0.498 rad/s. A Kanayama control law with gains 1.6, 3.2, and 2.4 was used. The controller always received the estimated state. True-state feedback was used as an implementation control and yielded a mean error of 0.030 m after the transient.

The true state started on the reference trajectory. The filter initial state was perturbed with a Gaussian distribution and diagonal covariance. The initial standard deviations were 1 m, 1 m, 15 degrees, 0.2 m/s, 0.1 rad/s, and 0.01 rad/s for the six components in (1).

B. Sensors

Four channels were simulated. GNSS measured position at 5 Hz. Odometry measured longitudinal velocity at 50 Hz. The gyroscope measured $\omega + b_g$ at 100 Hz. The magnetometer measured heading at 10 Hz. Nominal standard deviations were 0.60 m for each GNSS component, 0.05 m/s for odometry, 0.012 rad/s for the gyroscope, and 0.035 rad for the magnetometer.

The observation matrices were linear in the state. The state follows the definition in (1). The observation matrices for each channel were defined in (3).

$$\begin{aligned} \mathbf{H}_{\text{gnss}} &= \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \\ \mathbf{H}_{\text{odo}} &= \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \\ \mathbf{H}_{\text{gyro}} &= \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 1 \end{bmatrix} \\ \mathbf{H}_{\text{mag}} &= \begin{bmatrix} 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \end{aligned} \quad (3)$$

The magnetometer innovation is angular. Before computing the statistic, it was normalized to the interval $(-\pi, \pi]$ using the function $\text{atan2}(\sin(\cdot), \cos(\cdot))$ to avoid discontinuities around $\pm\pi$.

The nonlinearity of the problem was confined to propagation. This choice prevents differences from a visual-perception or LiDAR front end from dominating the comparison of the adaptive mechanisms.

C. Degradation Protocol

Each mode was parameterized by a severity s between zero and one. Severity zero corresponds to the nominal channel. Severity one reaches the maximum magnitude defined before running the estimators. Noise scaling, drifting bias, outliers, intermittent loss, and a combined scenario were studied.

For variable GNSS noise, the standard deviation scaled as $\sigma(s) = \sigma_0(1 + 9s)$, reaching a factor of ten at $s = 1$. Magnetometer bias grew as the linear ramp $b(t, s) = 8s \sigma_0 \min(t - t_0/T - t_0, 1)$, where $t_0 = 20$ s is the onset time and $T = 120$ s is the run horizon. GNSS outliers were generated with probability $p_{\text{out}} = 0.30$ s per sample and fixed magnitude $25 \sigma_0$. Loss episodes started with probability $p_{\text{drop}} = 0.05$ s per sample and had exponentially distributed duration with mean 40s channel samples. For the two-dimensional GNSS channel, each outlier component received an independently sampled sign of +1 or -1, yielding an additive perturbation of $\pm 25 \sigma_0$ independently along each coordinate.

For the pure data-loss scenario, the mean episode duration was defined as 40s GNSS channel samples, not seconds of physical time. Because GNSS is sampled at 5 Hz, severity one corresponds to a mean of 40 samples, or approximately 8 s.

The combined scenario simultaneously affected GNSS and the magnetometer. GNSS was subjected to variable noise with effective factor $(1 + 5.4s)$, outliers with probability 0.15s, and loss with probability 0.025s. The magnetometer was subjected to bias with the same ramp as in the pure-bias scenario. Severity s was common to both channels. The five evaluated values were $s \in \{0, 0.25, 0.50, 0.75, 1.00\}$.

In the combined scenario, the GNSS noise component used an effective severity of 0.6 s , whereas the outlier and dropout components used 0.5 s . Consequently, the noise standard-deviation factor was $1 + 5.4s$, the outlier probability was 0.15 s , the loss-entry probability was 0.025 s , and the mean loss duration was 20 s GNSS channel samples, where s denotes degradation severity.

The secondary exploratory fine-grid analysis used severity values of 0.02, 0.04, 0.07, 0.12, 0.18, and 0.25 for the bias and combined scenarios, with 40 repetitions per cell. It was used only to locate the consistency-loss boundary and was not used to tune any estimator.

Each run lasted 120 s. Degradation started at 20 s. Five severity levels were evaluated. Each cell contained forty repetitions. Noise realizations were shared by all estimators within a cell, enabling paired comparisons. The confirmatory evaluation used

new seeds to analyze bias compensation and replicate the common arms. Table I summarizes the protocol.

TABLE I
LOCKED PARAMETERS OF THE SIMULATION PROTOCOL

Parameter	Value	Unit
Run duration	120	s
Base frequency	100	Hz
Repetitions per cell	40	repetitions
Severity levels	0, 0.25, 0.50, 0.75, 1.00	dimensionless
Maximum noise factor	10	times the nominal standard deviation
Terminal bias	8	nominal standard deviations
Maximum outlier rate	0.30	per sample
Outlier magnitude	25	nominal standard deviations
Maximum loss-onset rate	0.05	per sample
Maximum mean loss duration	40	channel samples

D. Base Filter And Comparators

The common core was an extended Kalman filter with sequential channel-wise updates. Prediction used the Jacobian of the unicycle model, and covariance updating used the Joseph form. All estimators shared the same core and differed only in the covariance adapter or the action applied after detection.

The nominal EKF used covariances derived from the simulator. The process-noise scale q was selected through a one-dimensional sweep over sixty repetitions with seeds independent of evaluation, choosing the value that brought the expected NEES closest to six. The resulting value was $q = 0.9976$, with a mean NEES of 5.999. The mistuned EKF used $0.1q$, yielding an overconfident filter with NEES around twenty-four.

Covariance matching estimated a scalar factor by matching the trace of the sample innovation covariance over a thirty-sample window to the trace expected under the model [31]. For Sage-Husa [33], $d_k = \frac{1-b}{1-b^{k+1}}$, with forgetting factor $b = 0.98$, and the scalar multiplier was updated as $\lambda_k = (1 - d_k)\lambda_{k-1} + d_k \frac{\text{tr}(v_k v_k^T - HPH^T)}{\text{tr}(R_{\text{nom}})}$, followed by clipping to the interval $[0.2, 1000]$. Here, v_k is the channel innovation, HPH^T is the predicted measurement covariance, R_{nom} is the nominal measurement covariance, $\text{tr}(\cdot)$ denotes the trace, and the index k denotes successive updates of the corresponding sensor channel. Huber used the influence function when the standardized innovation norm was within the threshold and a bounded covariance multiplier otherwise, with $\kappa = 1.345$, implemented as iterative reweighting of the innovation [32]. These are channel-wise scalar versions and do not represent every matrix-valued variant of each family.

The scalar covariance-matching rule used (4) and (5), where L is the window length, v_j is the j th innovation in the window, HPH^T is the predicted measurement covariance, R_{nom} is the nominal measurement covariance, and R_{est} is the estimated

measurement covariance. The multiplier λ was clipped to the interval from 1 to 1000.

$$R_{\text{est}} = \frac{1}{L} \sum_{j=1}^L v_j v_j^T - HPH^T \quad (4)$$

$$\lambda = \text{clip}\left(\frac{\text{tr}(R_{\text{est}})}{\text{tr}(R_{\text{nom}})}, 1, 1000\right) \quad (5)$$

For the robust Huber comparator, the effective covariance multiplier λ was computed from the standardized innovation norm r using (6), where NIS is the normalized innovation squared of the channel, d is its dimension, κ is the Huber threshold, and $\lambda_{\text{max}} = 1000$ is the upper bound of the multiplier. This is the exact weight-equivalent implementation used in the simulations.

$$r = \sqrt{\text{NIS}/d},$$

$$\lambda = \begin{cases} 1 & r \leq \kappa \\ \min\left(\left(\frac{r}{\kappa}\right)^2, \lambda_{\text{max}}\right) & \text{otherwise} \end{cases} \quad (6)$$

Table II presents the parameters that distinguish the estimators.

TABLE II
REPRODUCIBLE CONFIGURATION OF THE ESTIMATORS

Estimator	MECHANISM	Parameters
Nominal EKF	Fixed R	Q scale $q = 0.9976$
Mistuned EKF	Fixed R	Q scale $0.1q = 0.09976$
Covariance matching	Trace-based scaling	window $L = 30$, $\lambda \in [1, 1000]$
Sage-Husa	Trace-based scaling	$b = 0.98$, $\lambda \in [0.2, 1000]$
Huber	Equivalent weighting	$\kappa = 1.345$, $\lambda \in [1, 1000]$
Two moments	Inflation and gating	$\rho = 0.5$, dead zone $\delta = 0.5$ NIS level 0.01, mean level 0.001, $\beta = 0.02$, three faults, dwell 25
Compensation	Augmented state	bias $P_0 = 0.01R_{i,0}$, density $0.1\sigma_{0,i}$, latched activation

Except for the two test levels, which prescribe false-alarm thresholds, the numerical settings in Table II were engineering choices fixed during exploratory development. The pilot informed the adoption of the first moment mechanism after slow bias was not detected by second moment methods. It was not used for formal hypothesis testing. Before the primary and confirmatory evaluations, all settings were locked and were not retuned. These evaluations used seed sets disjoint from the pilot. The dwell was counted in channel updates because the detection and readmission decisions are evaluated only when a measurement from that channel arrives. A common count therefore specifies a common number of decision opportunities, while its duration in seconds varies with the sampling rate. The stricter first moment level of 0.001 limits false exclusions because it governs channel removal. The 0.01 level governs second moment adaptation.

E. Two-Moment Adaptation

For each channel i with d_i degrees of freedom, measurement $\mathbf{z}_{i,k}$, and observation matrix $\mathbf{H}_i$, the innovation with respect to the predicted state $\hat{\mathbf{x}}_k^-$ was

$$\mathbf{v}_{i,k} = \mathbf{z}_{i,k} - \mathbf{H}_i \hat{\mathbf{x}}_k^- \quad (7)$$

The statistic was computed using the current inflation factor $\lambda_{i,k-1}$. With the predicted state covariance $\mathbf{P}_k^-$ and the nominal measurement covariance $\mathbf{R}_{i,0}$ of channel i , its covariance was

$$\mathbf{S}_{i,k} = \mathbf{H}_i \mathbf{P}_k^- \mathbf{H}_i^T + \lambda_{i,k-1} \mathbf{R}_{i,0} \quad (8)$$

The second moment and its ratio to the expected value were defined as

$$\begin{aligned} \epsilon_{i,k} &= \mathbf{v}_{i,k}^T \mathbf{S}_{i,k}^{-1} \mathbf{v}_{i,k}, \\ q_{i,k} &= \frac{\epsilon_{i,k}}{d_i} \end{aligned} \quad (9)$$

Inflation used a dead zone $\delta = 0.5$, an exponent $\rho = 0.5$, and an upper bound $\lambda_{max} = 1000$.

$$\lambda_{i,k} = \begin{cases} \min(\lambda_{i,k-1} q_{i,k}^\rho, \lambda_{max}) & q_{i,k} > 1 + \delta \\ \max(\lambda_{i,k-1} q_{i,k}^\rho, 1) & q_{i,k} < 1 - \delta \\ \lambda_{i,k-1} & \text{otherwise} \end{cases} \quad (10)$$

The covariance applied in the update was given by (11).

$$\mathbf{R}_{i,k} = \lambda_{i,k} \mathbf{R}_{i,0} \quad (11)$$

The first moment was computed from the Cholesky factor $\mathbf{L}_{i,k}$ of $\mathbf{S}_{i,k}$ in (8), where $\mathbf{r}_{i,k}$ is the whitened innovation, $\mathbf{m}_{i,k}$ is its exponentially weighted mean, and β is the smoothing factor.

$$\mathbf{r}_{i,k} = \mathbf{L}_{i,k}^{-1} \mathbf{v}_{i,k} \quad (12)$$

$$\mathbf{m}_{i,k} = (1 - \beta) \mathbf{m}_{i,k-1} + \beta \mathbf{r}_{i,k}, \quad \beta = 0.02 \quad (13)$$

Under the nominal hypothesis, the asymptotic covariance of this mean is $\beta/(2 - \beta)$ times the identity. The shift statistic was

$$T_{i,k} = \frac{2 - \beta}{\beta} \mathbf{m}_{i,k}^T \mathbf{m}_{i,k} \quad (14)$$

The gate declared a fault when $\epsilon_{i,k}$ exceeded the chi-square quantile with probability 0.99 and d_i degrees of freedom, or when $T_{i,k}$ exceeded the 0.999 quantile with the same dimension. Three consecutive faults excluded the channel. Exclusion lasted twenty-five updates of that channel. This corresponds to physical dwell times of 5 s for GNSS (5 Hz), 0.5 s for odometry (50 Hz), 0.25 s for the gyroscope (100 Hz), and 2.5 s for the magnetometer (10 Hz). Readmission required the dwell counter to reach zero and the test to produce no fault at the next update.

During exclusion, $\lambda_{i,k}$ and $\mathbf{m}_{i,k}$ were held unchanged and the fault counter was reset to zero.

Algorithm 1. Two-moment adaptation for channel i at step k

1. Compute the innovation $\mathbf{v}_{i,k}$ using (7). For the magnetometer channel, normalize it to $(-\pi, \pi]$.
2. Compute covariance $\mathbf{S}_{i,k}$ using the current $\lambda_{i,k-1}$, according to (8).
3. Compute $\epsilon_{i,k}$ and $q_{i,k}$ according to (9). Update $\lambda_{i,k}$ using (10).
4. Whiten the innovation using $\mathbf{L}_{i,k}$ and update the mean $\mathbf{m}_{i,k}$ according to (12) and (13). Compute $T_{i,k}$ according to (14).
5. If $\epsilon_{i,k}$ exceeds the 99% chi-square threshold, or $T_{i,k}$ exceeds the 99.9% threshold, increment the consecutive-fault counter. Otherwise, reset it to zero.
6. If the counter reaches three while the channel is active, start exclusion with the dwell counter set to twenty-five. While the channel remains excluded, skip Steps 3 and 4 and decrement the dwell counter.
7. When the dwell counter reaches zero, evaluate the test at that same step. If no fault is generated, readmit the channel. If the test still generates a fault, keep the channel excluded with the counter at zero and reevaluate at every subsequent step until the first fault-free update.
8. If the channel remains active, perform the Joseph-form update.

The proposed estimator used exclusion as the primary action. The secondary compensation action expanded the state from six to eight components by adding two additive-bias states, one for the magnetometer and one for odometry. These states were initialized with variance $0.01 R_{i,0}$ for each channel and zero process noise before detection, so they remained inert and did not distort nominal estimation. After three consecutive first-moment faults, the process noise of the corresponding state was released with density $0.1 \sigma_{0,i}$ and activation was latched, meaning that it was not deactivated even if the channel was readmitted. Permanent compensation used the same structure with activation from the third update, without requiring a trigger. No GNSS bias state was added because a constant position shift cannot be separated from the position state without another absolute position reference to constrain it.

F. Metrics And Statistical Analysis

State-wise RMSE, position RMSE, NEES, channel-wise NIS, tracking RMSE, maximum excursion, control effort, recovery time, and computational cost were computed. Recovery time was defined as the elapsed time from degradation onset at 20 s until the last instant at which the tracking error of that repetition and estimator exceeded a band equal to twice the pre-degradation median of the same estimator in the same repetition, computed over the first twenty seconds. Runs whose error did not return to the band were assigned the remaining simulation horizon, 99.99 s, and treated as truncated observations. A complementary sensitivity analysis required

the error to remain inside the band during the final five seconds of the run and summarized the fraction of repetitions without recovery. For H2, inconsistency was quantified as $I = |\log(\frac{NEES}{6})|$, so that departures above and below the expected six-dimensional NEES value were treated symmetrically on a relative scale.

Comparisons were paired by repetition within each scenario-severity cell. The Wilcoxon signed-rank test was applied. For H3, Holm correction was applied to the thirty available switching comparisons (three ablation scenarios, five severities, and two channels). For the paired comparison at maximum severity, Holm correction was applied jointly to the ninety comparisons formed by crossing the six scenarios, five comparators, and three reported performance metrics. These metrics were position RMSE, tracking RMSE, and maximum excursion. Rankings were evaluated with Spearman correlations computed over the six main estimators within each cell. Bimodal distributions were summarized by the median, interquartile range, coefficient of variation, and fraction of repetitions showing improvement. The nominal cost of each mechanism was computed for each repetition as the percentage increase in nominal-condition position RMSE relative to the paired nominal EKF in the same repetition, and the median of the forty individual increases was reported.

Inference is comparative and predictive within the simulator. The primary evaluation included 240 combinations, consisting of 180 main-estimator combinations and 60 ablation combinations, with forty repetitions each, for a total of 9,600 observations. The confirmatory evaluation replicated those 240 combinations with disjoint seeds and added 60 compensation-arm combinations, yielding 300 combinations and 12,000 observations. The six main estimators covered all conditions. The four ablations were restricted to the nominal, bias, and outlier conditions.

Consistency was evaluated against a two-sided 95% chi-square band. For NEES, the forty repetitions were averaged before constructing the band for a six-dimensional state. For NIS, the degrees of freedom of each channel were retained. A value inside the band indicates that consistency was not rejected under the simulated model. It does not demonstrate that the complete error distribution is correct or that the filter is safe for every controller.

The main comparisons included each estimator against the proposed scheme in all six conditions at maximum severity. Counts of wins, losses, and no significant difference were obtained over the thirty comparisons for each metric and did not depend on selecting only significant results.

The ranking comparison used arm means within each cell. The target metrics were tracking RMSE, maximum excursion, and recovery time. They represented average behavior, peak behavior, and persistence of the transient response, respectively.

G. Implementation Controls And Falsification

Sampling rates and dispersion were checked against their analytical values. The nominal EKF was calibrated only under

no degradation and achieved a NEES of 5.999. NIS values approached the degrees of freedom of the four channels. The controller was verified with the true state, and a negative control was constructed with a linearized plant and a correctly specified filter.

The negative control evaluated eight fixed covariance scales. In the degraded linear plant, the rankings by estimation and tracking error had a Spearman correlation of 0.976. In the nonlinear plant, the nominal and degraded conditions produced a correlation of 1.000. These results supported the implementation of the H1 metrics.

The main execution used batched linear algebra to advance the forty repetitions in parallel. A scalar implementation processed the repetitions separately. In six cases, the confidence intervals of the means overlapped and the ratios between coefficients of variation ranged from 0.75 to 1.32. This is an internal verification rather than a reproduction by a second team.

The development, primary-evaluation, and confirmatory-evaluation sets used disjoint seeds. Parameters were fixed before each evaluation and were not estimated from its realizations.

III. RESULTS

The primary evaluation, with 9,600 observations and seeds independent of calibration, supports the consistency results in Section 3A, the H1 and H2 analyses in Section 3B, the ablation study in Section 3D, and the paired comparisons in Section 3E. The secondary fine-grid analysis supports the consistency-boundary statement in Section 3A. The confirmatory evaluation with disjoint seeds, with 12,000 observations, supports Table III, Table IV, and the numerical results in Sections 3C and 3F. Section 3G contrasts the two sets.

A. Consistency Across Severity

The nominal EKF remained consistent under no degradation and data loss. It became inconsistent under noise, outliers, bias, and the combined scenario. Covariance matching, Sage-Husa, and robust Huber weighting kept NEES within the theoretical band under noise and outliers. None remained consistent under drifting bias.

Table III summarizes position RMSE at maximum severity. Values are the median and interquartile range in meters.

TABLE III
POSITION RMSE BY ESTIMATOR AND SCENARIO

Estimator	No degraded n	Noise	Bias	Outliers	Data loss	Combi ned
Nominal	0.059	0.415	0.732	0.573	0.079	0.896
EKF	[0.052, 0.067]	[0.379, 0.510]	[0.728, 0.740]	[0.516, 0.621]	[0.062, 0.098]	[0.844, 0.965]
Mistuned	0.072	0.395	0.720	0.527	0.088	0.874
EKF	[0.061, 0.084]	[0.347, 0.486]	[0.713, 0.729]	[0.496, 0.605]	[0.073, 0.110]	[0.810, 0.923]
Covariance	0.059	0.116	0.744	0.096	0.076	0.763
matching	[0.054, 0.069]	[0.084, 0.138]	[0.740, 0.754]	[0.070, 0.128]	[0.061, 0.101]	[0.754, 0.769]
Sage-	0.061	0.130	0.750	0.098	0.081	0.766
Husa	[0.054, 0.071]	[0.088, 0.146]	[0.744, 0.754]	[0.075, 0.130]	[0.068, 0.103]	[0.755, 0.772]
Robust	0.065	0.110	0.746	0.073	0.084	0.761
Huber	[0.058, 0.074]	[0.093, 0.163]	[0.741, 0.754]	[0.063, 0.087]	[0.063, 0.101]	[0.755, 0.769]
Triggered	0.072	0.124	0.187	0.093	0.084	0.745
exclusion	[0.061, 0.082]	[0.084, 0.179]	[0.122, 0.747]	[0.074, 0.135]	[0.073, 0.102]	[0.294, 0.763]

Fig. 1 and Fig. 2 show NEES as a function of severity. On the fine grid for the bias scenario, all adaptive estimators lost consistency at severity 0.04. The two-moment scheme did not shift this boundary. Its effect was to attenuate the severity of the inconsistency after the threshold. At severity 0.25, it produced a NEES of 75.3 versus 84.1 for the best comparator. At severity one, the difference increased, although the filter remained inconsistent.

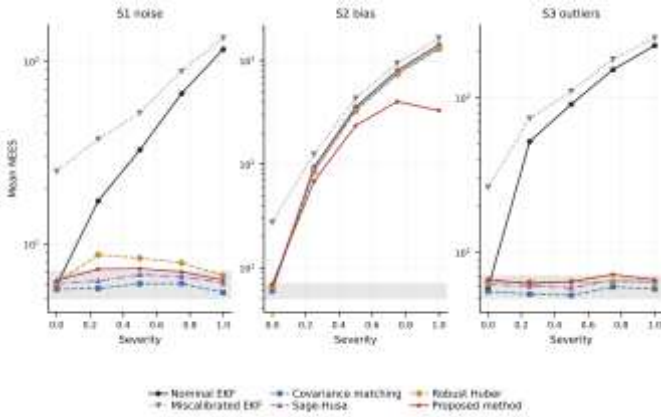


Fig. 1. Mean NEES (average of forty repetitions) as a function of severity for S1 GNSS noise, S2 magnetometer bias, and S3 GNSS outliers. The gray shaded band corresponds to the two-sided 95% consistency interval for a six-dimensional state with forty repetitions. The vertical scale is logarithmic.

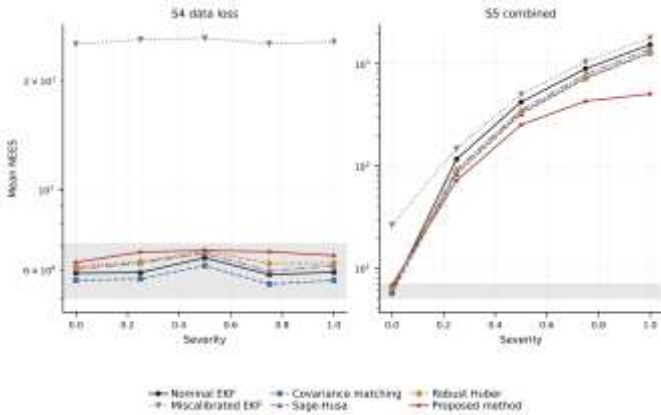


Fig. 2. Mean NEES (average of forty repetitions) as a function of severity for the S4 GNSS data-loss scenario and the combined S5 scenario. The shaded band and vertical scale follow the same definition as in Fig. 1.

B. Transfer From Estimation To Control

The ranking of the six estimators by position RMSE was almost completely preserved when they were ranked by tracking RMSE. The mean Spearman correlation was 0.994 across thirty scenario-severity combinations. The minimum was 0.943. H1 was rejected for average performance.

The relationship was weaker for transient metrics. The mean correlation was 0.899 for maximum excursion and 0.491 for recovery time. Some cells showed negative correlation for recovery. After removing the nominal scenario and all zero-severity conditions, the correlations were 0.991, 0.897, and 0.523. The conclusion therefore did not depend on repeated nominal conditions.

The sensitivity analysis requiring the final five seconds to remain inside the band yielded a mean correlation of 0.401 between RMSE and the unrecovered fraction across twenty-nine informative cells. Under degraded conditions, it was 0.546 across nineteen cells. One cell did not admit a correlation because all estimators had the same fraction. This sensitivity analysis confirmed that average RMSE contains less information about transient behavior.

Inconsistency, quantified as $|\log(\frac{NEES}{6})|$, correlated less strongly with tracking error than RMSE. The mean absolute correlations were 0.728 and 0.994. The paired difference was statistically significant according to the Wilcoxon signed-rank test, with $p < 0.0001$. H2 was rejected. Consistency was useful as an adaptation signal but was not the best criterion for selecting the estimator with the lowest mean control error. Fig. 3 summarizes these correlations and shows their global distribution and the cell-wise pattern for recovery time.

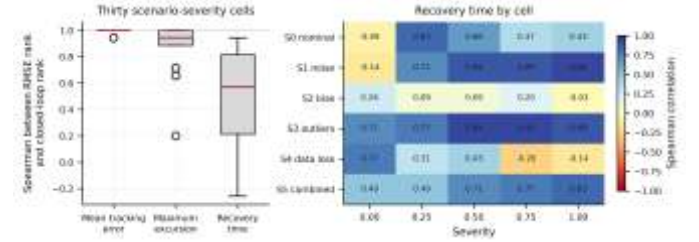


Fig. 3. Left: box plot of the Spearman correlation between the ranking of the six estimators by estimator position RMSE and their ranking by three closed-loop metrics, computed over thirty scenario-severity cells with forty repetitions each. Right: heat map of the correlation for recovery time, broken down by cell. Values inside the cells are point correlations.

C. Blind Spot Under Drifting Bias

In the bias scenario, the five comparators gradually followed the shifted measurement. The instantaneous innovation remained moderate, and energy-based adaptation did not identify the mean shift. At severity one, median position RMSE ranged from 0.720 to 0.750 m for filters without the first-moment test.

The first-moment test produced a marked reduction when it triggered. In the evaluation with disjoint seeds, exclusion reached a median of 0.187 m versus 0.732 m for the nominal EKF. However, the third quartile was 0.747 m. The mechanism

improved 28 of 40 repetitions and provided no benefit in the remaining twelve. Its coefficient of variation was 0.789, compared with values near 0.01 for the comparators. The distribution was bimodal.

The ablation without the first-moment test reproduced the behavior of covariance matching, Sage-Husa, and robust Huber weighting. This observation identifies mean monitoring as the element responsible for the response to bias. It does not show that the proposal detects smaller biases. The fine grid showed the same consistency-loss boundary for all adaptive estimators. Under the final-dwell sensitivity criterion, the comparators failed to recover in every repetition, whereas the proposed scheme failed to recover in 45% of repetitions.

Fig. 4 presents medians and interquartile ranges. The wide band of the proposed scheme at high severity reflects the mixture of trigger and non-trigger runs.

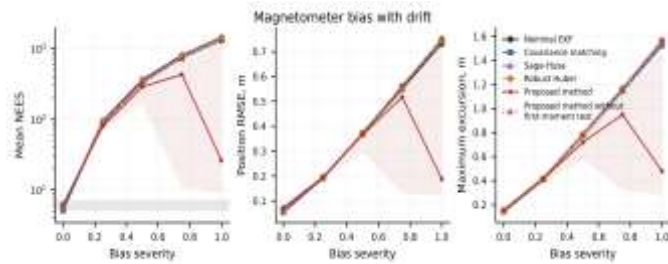


Fig. 4. Position RMSE in meters (left), NEES (center), and maximum excursion in meters (right) as functions of severity for the S2 magnetometer-bias scenario. Forty repetitions per cell. Lines show medians and shaded bands show interquartile ranges. The red dotted line represents the ablation without the first-moment test. NEES is displayed on a logarithmic scale with the consistency band shaded.

D. Dwell Time And Ablations

The ablation without minimum dwell time reacted more rapidly to bias but generated false alarms under nominal conditions. The complete scheme produced 3.70 switches per run in GNSS and 1.30 in the magnetometer. The ablation produced 12.63 and 18.59, respectively, corresponding to a reduction factor between 3.4 and 14.3. All thirty paired comparisons remained significant after Holm correction.

Dwell time has a sensitivity cost that should be reported. In the primary evaluation, the no-dwell ablation obtained a median position RMSE of 0.141 m in the bias scenario at severity one, versus 0.149 m for the complete scheme, because it triggers earlier at the cost of more false alarms. Under nominal conditions, its RMSE increased to 0.087 m versus 0.060 m for the complete scheme, and its NEES moved away from six because of spurious exclusions in the absence of degradation. Fig. 5 shows that dwell time trades bias sensitivity for fewer false alarms under nominal conditions. The dead zone provided a smaller but consistent benefit in both scenarios. H3 was supported for reducing switching, with the sensitivity trade-off quantified.

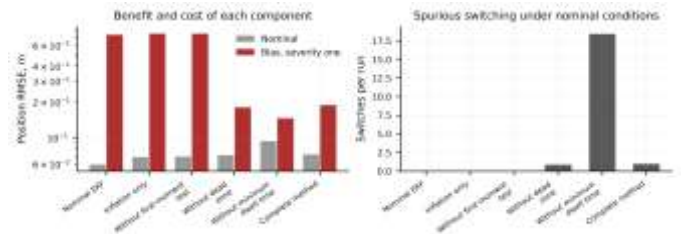


Fig. 5. Top: position RMSE in meters under the nominal scenario (left) and bias scenario (right) for the proposed scheme and its four ablations, showing the median and interquartile range across forty repetitions. Bottom: mean number of switches between accept and exclude under nominal conditions, by channel, for the complete scheme and the no-minimum-dwell-time ablation.

E. Overall Comparison And Cost

At severity one, the exclusion scheme won fourteen of thirty position-RMSE comparisons, lost three, and showed no significant difference in thirteen. For tracking RMSE, it won fifteen, lost five, and tied ten. For maximum excursion, it won fifteen, lost one, and tied fourteen. There was no uniform superiority.

The relative computational cost of the vectorized implementation was 1.26 with respect to the nominal EKF. Robust Huber weighting cost 1.03, Sage-Husa 1.10, and covariance matching 1.19. Measurements were obtained using Python 3.12.10 with NumPy 2.5.2 and SciPy 1.18.1 on an Intel Core i5-10400T processor at 2.0 GHz. These values are ratios within the same implementation and do not represent execution times on an embedded processor. Fig. 6 presents cost and paired comparisons.

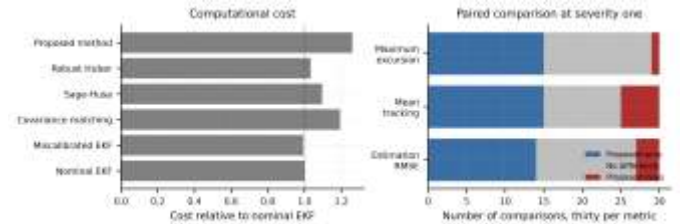


Fig. 6. Left: computational cost of each estimator relative to the nominal EKF in the same implementation. Right: Wilcoxon paired-comparison outcomes between the proposed scheme and each comparator at severity one for three closed-loop metrics. Grouped colored bars show wins, ties, and losses of the proposed scheme over the thirty scenario-comparator pairs.

F. Exclude Or Compensate

The confirmatory evaluation compared exclusion, triggered compensation, and permanent compensation. Table IV shows that no mechanism dominated all three objectives.

TABLE IV
TRADE-OFF BETWEEN EXCLUSION AND COMPENSATION UNDER BIAS

Mechanism	No degradation	Bias, severity one	Reduction vs. nominal	Repetitions improved	Coefficient of variation
Nominal EKF	0.059 [0.052, 0.067]	0.732 [0.728, 0.740]	Reference	Reference	0.012
Triggered exclusion	0.072 [0.061, 0.082]	0.187 [0.122, 0.243]	74.5 %	28 of 40	0.789
Triggered compensation	0.077 [0.070, 0.086]	0.243 [0.208, 0.303]	67.0 %	40 of 40	0.555
Permanent compensation	0.088 [0.079, 0.106]	0.200 [0.187, 0.221]	72.9 %	40 of 40	0.143

Exclusion yielded the best median under bias, 0.187 m, with a nominal cost of 15.9%. Its response was bimodal and it improved only 28 of 40 repetitions. Triggered compensation yielded a median of 0.243 m and improved all forty repetitions. Its nominal cost was 32.2%. Permanent compensation yielded a median of 0.200 m and the lowest dispersion, but increased nominal error by 53.5%.

Fig. 7 represents this trade-off. Triggered compensation reduced the risk of inaction but did not combine the advantages of the other strategies at no cost. The choice depends on whether the priority is the median, worst-case behavior, or nominal performance.

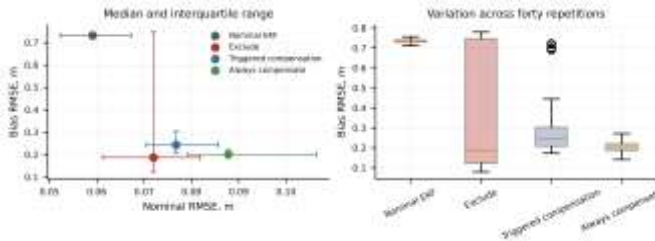


Fig. 7. Trade-off among the three bias-response mechanisms evaluated in the second confirmatory round with forty repetitions and disjoint seeds. The horizontal axis shows position RMSE in meters under nominal conditions, and the vertical axis shows position RMSE under the bias scenario at severity one. Each point represents the median of one mechanism. Error bars correspond to the interquartile range.

G. Replication With Disjoint Seeds And Numerical Verification

The ten common arms were evaluated using two disjoint seed sets. Median relative differences between sets ranged from 0.5% to 5.9% in magnitude. Cell-wise maxima reached 27% because realizations shared within each set shifted together. The qualitative conclusions were preserved. The fractions of improvement for the exclusion scheme were 0.775 and 0.700. Median improvements were 79.7% and 74.5%.

The scalar implementation reproduced six cases. All confidence intervals of the means overlapped. Ratios between coefficients of variation ranged from 0.75 to 1.32. For exclusion under bias, the coefficients were 0.88 in the batched execution and 0.91 in the scalar execution. This verification reproduced the

bimodality, although it does not constitute an implementation by a second team.

IV. DISCUSSION

A. Main Findings

On the studied platform, estimation RMSE almost completely preserved the ranking of mean tracking error. H1 was not supported either in the negative control or in the six-estimator benchmark. The result supports RMSE for comparing average performance within similar systems, although not as the only metric. The association decreased for maximum excursion and, more clearly, for recovery. The final-dwell sensitivity analysis led to the same conclusion.

H2 was not supported either. Statistical consistency did not predict mean control error better than RMSE. NEES remains valuable for interpreting uncertainty and governing adaptation, but it does not directly represent the cost experienced by the controller.

Mechanisms based only on innovation energy responded to variable noise and outliers but not to a slow mean shift. The first-moment test added complementary information. The ablation attributed the observed response to bias to this test rather than to bounded inflation alone. However, exclusion failed in twelve of forty repetitions and did not shift the consistency boundary.

H3 was supported for reducing spurious switching, at the expense of lower bias sensitivity. Post-detection actions occupied different points on a trade-off frontier. Exclusion had the lowest nominal cost and a favorable median but exhibited a bimodal response. Triggered compensation reduced the risk of inaction at the cost of higher nominal error. Permanent compensation was more stable under bias and incurred the largest penalty under no degradation.

B. Relationship To Previous Work

The results agree with literature using NIS to adapt covariances [2,3]. They are also compatible with strategies that separate time scales [4]. The contribution is not to present adaptive NIS or sequential mean monitoring as new concepts [2,29,30]. Rather, it lies in their channel-wise integration, the demonstration of a blind spot under mean drift, and the comparison of post-detection actions under a common protocol.

The response of covariance matching, Sage-Husa, and robust Huber weighting to noise and outliers confirms that the second moment is appropriate when dispersion changes. Their common failure under bias shows that robustness to outliers is not equivalent to tolerance of slow drift. This distinction is relevant to systems that combine absolute and relative sensors.

Mobile-localization studies under degraded GPS have shown benefits from adaptation [7,8]. The present study adds a closed-loop evaluation and continuous severity parameterization. However, it does not compare perception front ends or real data. It therefore complements rather than replaces those studies.

C. Design Implications

An adaptive estimator should recognize variance and mean changes separately. It must then decide whether to down-weight, exclude, or compensate. A single inflation factor does not perform all three functions.

Dwell time depends on the cost of false alarms. Reducing it increases sensitivity and nominal switching. When false exclusion is costly, dwell time provides protection. If immediate response is required, more false alarms can be accepted or an active bias state can be maintained.

Nominal cost should be reported alongside robustness. Reporting only degraded conditions favors mechanisms that keep additional states active. The comparison showed that this cost can exceed 50% even with effective compensation.

Distribution shape should accompany the average result. Exclusion produced two groups. When the test triggered, error dropped markedly. When it did not, error remained near that of the nominal filter. The median, interquartile range, and improvement fraction described this structure better than a global mean.

A method can have a favorable RMSE while retaining a non-negligible probability of inaction. In safety-related applications, the success fraction and worst quartile may be more informative than the mean. Triggered compensation had a worse median than exclusion but eliminated no-improvement runs in this experiment.

Position RMSE predicted tracking RMSE well because both aggregate error over the same window. Recovery depends on the temporal sequence. A filter can accumulate little error yet take longer to return to the band. Therefore, validation for control requires both an average metric and a transient metric.

D. Limitations

The evidence comes from one simulated plant and one control law. Observations were linear in the state. Real slip, changing friction, saturation, visual perception, LiDAR, and electromagnetic interference were not represented. The synthetic protocol enabled repetition and controlled severity, but its distributions were not fitted to physical sensors. It also did not model dependencies such as temperature-drift or latency-packet-loss relationships.

The first-moment test did not shift the consistency boundary and its triggering was unreliable for exclusion. The chi-square thresholds assume Gaussian noise, although outliers violate that assumption. The false-alarm rate under heavy-tailed noise was not explicitly controlled, nor were thresholds optimized using a control cost. Compensation is a secondary analysis evaluated with disjoint seeds and requires confirmation in an evaluation specified before execution.

Compensation states were added to the magnetometer and odometry. A constant GNSS bias was not observable without another absolute reference. Therefore, conclusions about compensation are limited to channels in which dynamic or

sensor redundancy allows the bias to be separated from the physical state.

Computational cost comes from a vectorized Python implementation. The ratios compare methods in the same environment but do not demonstrate real-time execution on an embedded processor. Reproduction by another team and validation with physical sensors are also required before extrapolating performance to a real platform.

V. CONCLUSION

The results demonstrate that adaptive multisensor fusion should not be evaluated solely through estimation accuracy, because average and transient closed-loop behavior may convey different information. In the studied platform, position RMSE was an excellent predictor of mean tracking performance, with a mean Spearman correlation of 0.994, but its association with recovery time decreased to 0.491, showing that transient response requires separate evaluation. Methods based primarily on innovation energy successfully adapted to changes in measurement dispersion, such as increased noise and outliers, but remained unable to address slowly varying bias. Incorporating the first moment of the whitened innovation provided complementary sensitivity to persistent mean shifts and enabled differentiated post-detection actions. Unlike covariance matching, Sage-Husa estimation, and Huber weighting, which only modify the weight of a measurement according to its dispersion, the proposed scheme monitored persistent mean shifts separately from dispersion changes and coupled their detection with an explicit exclusion or compensation action. However, the triggered-exclusion strategy remained bimodal, improving 28 of 40 repetitions under maximum bias severity, whereas triggered bias compensation improved all 40 repetitions and achieved a median reduction of 67.0% relative to the nominal EKF, at the cost of a 32.2% increase in nominal error. Minimum dwell time substantially reduced spurious switching, although this robustness was accompanied by lower sensitivity to bias. Taken together, these findings show that covariance adaptation, sensor exclusion, and bias compensation represent different operating trade-offs rather than interchangeable solutions, and that no single mechanism simultaneously maximizes nominal accuracy, robustness, response reliability, and transient performance. Future work should therefore evaluate the proposed framework with physical sensors, more realistic and coupled degradation mechanisms, and real-time implementation on embedded hardware to determine its practical robustness and generalizability beyond the simulated conditions considered here.

DATA AVAILABILITY STATEMENT

All files required to reproduce the study, including the source code, configuration files, random seeds, and scripts for generating the analyses, tables, and figures, are available in the project repository on GitHub at <https://github.com/fjburgosf/adaptive-two-moment-sensor-fusion>. The study uses a fully simulated platform and does not require restricted experimental data.

REFERENCES

- [1] Z. Chen, H. Biggie, N. Ahmed, S. Julier and C. Heckman, "Kalman Filter Auto-Tuning With Consistent and Robust Bayesian Optimization," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 60, no. 2, pp. 2236–2250, 2024., doi:10.1109/TAES.2024.3350587
- [2] L. Zeng, Z. Fan, G. He, N. Li, H. Chang, X. Yu and G. Yuan, "JC-AKF: Joint calibration and NIS-guided adaptive Kalman filtering for a low-cost MEMS IMU array," *Measurement*, early access, Art. no. 122475, 2026., doi:10.1016/j.measurement.2026.122475
- [3] M. A. J. Rabby, R. Kannan and K. Bingi, "Adaptive NIS-Guided Extended Kalman Filter for Robust State-of-Charge Estimation of Lithium-Ion Batteries," *Journal of The Electrochemical Society*, vol. 173, no. 12, Art. no. 120507, 2026., doi:10.1149/1945-7111/ae7347
- [4] B. Hang, W. Li, X. Li, R. Fan, H. Sun and F. Shen, "QA-MSCKF: A Statistically Adaptive Measurement Update Method for Visual-Inertial Odometry," *IEEE Internet of Things Journal*, vol. 13, no. 17, pp. 38934–38956, 2026., doi:10.1109/JIOT.2026.3704077
- [5] W. Hilal, N. Alsadi, S. A. Gadsden and M. A. AlShabi, "An adaptive SIF and KF estimation strategy for fault detection based on the NIS metric," in *Proc. Sensors and Systems for Space Applications XVI*, p. 26, 2023., doi:10.1117/12.2664054
- [6] X. Lai, G. Zhu and J. Chambers, "A fuzzy adaptive extended Kalman filter exploiting the Student's t distribution for mobile robot tracking," *Measurement Science and Technology*, vol. 32, no. 10, Art. no. 105017, 2021., doi:10.1088/1361-6501/ac0ca9
- [7] Y. Yu, B. Liu, G. Xie, A. Han, X. Huang and H. Deng, "Adaptive Kalman Filter-Based Multi-Sensor Fusion for Robust Localization in GPS-Degraded Environments," in *Proc. 2025 10th International Conference on Robotics and Automation Engineering*, pp. 144–149, 2025., doi:10.1109/ICRAE67496.2025.00030
- [8] W. Huang, L. Xiang, R. Chen, S. Xu and Q. Wang, "Adaptive Multi-Sensor Fusion Localization with Eigenvalue-Based Degradation Detection for Mobile Robots," *Sensors*, vol. 26, no. 5, Art. no. 1653, 2026., doi:10.3390/s26051653
- [9] Y. Geng and J. Wang, "Adaptive estimation of multiple fading factors in Kalman filter for navigation applications," *GPS Solutions*, vol. 12, no. 4, pp. 273–279, 2008., doi:10.1007/s10291-007-0084-6
- [10] F. Deng, H. Yang and L. Wang, "Adaptive Unscented Kalman Filter Based Estimation and Filtering for Dynamic Positioning with Model Uncertainties," *International Journal of Control, Automation and Systems*, vol. 17, no. 3, pp. 667–678, 2019., doi:10.1007/s12555-018-9503-4
- [11] Z. Chen, C. Heckman, S. Julier and N. Ahmed, "Weak in the NEES?: Auto-Tuning Kalman Filters with Bayesian Optimization," in *Proc. 2018 21st International Conference on Information Fusion*, pp. 1072–1079, 2018., doi:10.23919/ICIF.2018.8454982
- [12] L. Cai, B. Boyacioglu, S. E. Webster, L. van Uffelen and K. Morgansen, "Towards Auto-tuning of Kalman Filters for Underwater Gliders based on Consistency Metrics," in *Proc. OCEANS 2019 MTS/IEEE SEATTLE*, 2019., doi:10.23919/OCEANS40490.2019.8962573
- [13] G. LaMountain, J. Vilà-Valls and P. Closas, "Measurement noise covariance estimation in Gaussian filters: an online Bayesian solution," *EURASIP Journal on Advances in Signal Processing*, vol. 2025, no. 1, 2025., doi:10.1186/s13634-025-01215-w
- [14] W. Li, X. Lin, S. Li, J. Ye, C. Yao and C. Chen, "Robust autocovariance least-squares noise covariance estimation algorithm," *Measurement*, vol. 187, Art. no. 110331, 2022., doi:10.1016/j.measurement.2021.110331
- [15] B. Sun, Z. Zhang, D. Qiao, X. Mu and X. Hu, "An Improved Innovation Adaptive Kalman Filter for Integrated INS/GPS Navigation," *Sustainability*, vol. 14, no. 18, Art. no. 11230, 2022., doi:10.3390/su141811230
- [16] Z. Li, H. Zhang, Q. Zhou and H. Che, "An Adaptive Low-Cost INS/GNSS Tightly-Coupled Integration Architecture Based on Redundant Measurement Noise Covariance Estimation," *Sensors*, vol. 17, no. 9, Art. no. 2032, 2017., doi:10.3390/s17092032
- [17] R. Xin, J. Lin, S. Shi, R. Zhang, J. Zhang and J. Zhu, "Research on dual-threshold detection based adaptive fault-tolerant kalman filtering algorithm for dynamic 6-DOF measurement," *Mechanical Systems and Signal Processing*, vol. 210, Art. no. 111190, 2024., doi:10.1016/j.ymssp.2024.111190
- [18] C. Hajiyev, "Adaptive Filtering Against Sensor/Actuator Faults," *IFAC-PapersOnLine*, vol. 58, no. 4, pp. 336–341, 2024., doi:10.1016/j.ifacol.2024.07.240
- [19] D. Wu, L. Xia and J. Geng, "Heading Estimation for Pedestrian Dead Reckoning Based on Robust Adaptive Kalman Filtering," *Sensors*, vol. 18, no. 6, Art. no. 1970, 2018., doi:10.3390/s18061970
- [20] S. Liu, S. Li, J. Zheng, Q. Fu and Y. Yuan, "C/N0 Estimator Based on the Adaptive Strong Tracking Kalman Filter for GNSS Vector Receivers," *Sensors*, vol. 20, no. 3, Art. no. 739, 2020., doi:10.3390/s20030739
- [21] Z. Wang, B. Li, Z. Dan, H. Wang and K. Fang, "3D LiDAR Aided GNSS/INS Integration Fault Detection, Localization and Integrity Assessment in Urban Canyons," *Remote Sensing*, vol. 14, no. 18, Art. no. 4641, 2022., doi:10.3390/rs14184641
- [22] W. Wang, W. Shangguan, J. Liu and J. Chen, "Enhanced Fault Detection for GNSS/INS Integration Using Maximum Correntropy Filter and Local Outlier Factor," *IEEE Transactions on Intelligent Vehicles*, vol. 9, no. 1, pp. 2077–2093, 2024., doi:10.1109/TIV.2023.3312654
- [23] L. Qian, F. Qin, A. Li, K. Li and J. Zhu, "An INS/DVL integrated navigation filtering method against complex underwater environment," *Ocean Engineering*, vol. 278, Art. no. 114398, 2023., doi:10.1016/j.oceaneng.2023.114398
- [24] X. Wu, Z. Su, L. Li and Z. Bai, "Improved Adaptive Federated Kalman Filtering for INS/GNSS/VNS Integrated Navigation Algorithm," *Applied Sciences*, vol. 13, no. 9, Art. no. 5790, 2023., doi:10.3390/app13095790
- [25] Z. Mahmoudi, K. Nørgaard, N. K. Poulsen, H. Madsen and J. B. Jørgensen, "Fault and meal detection by redundant continuous glucose monitors and the unscented Kalman filter," *Biomedical Signal Processing and Control*, vol. 38, pp. 86–99, 2017., doi:10.1016/j.bspc.2017.05.004
- [26] G. Gao, B. Gao, S. Gao, G. Hu and Y. Zhong, "A Hypothesis Test-Constrained Robust Kalman Filter for INS/GNSS Integration With Abnormal Measurement," *IEEE Transactions on Vehicular Technology*, vol. 72, no. 2, pp. 1662–1673, 2023., doi:10.1109/TVT.2022.3209091
- [27] Y. Cheng, S. Zhang, X. Wang, H. Wang and H. Yang, "Kalman Filter with Adaptive Covariance Estimation for Carrier Tracking under Weak Signals and Dynamic Conditions," *Electronics*, vol. 13, no. 7, Art. no. 1288, 2024., doi:10.3390/electronics13071288
- [28] H. Kim, A. Bienkowski and K. R. Pattipati, "A Single-pass Noise Covariance Estimation Algorithm in Multiple-model Adaptive Kalman Filtering," in *Proc. 2023 IEEE Aerospace Conference*, pp. 1–9, 2023., doi:10.1109/AERO55745.2023.10115725
- [29] D. C. Lowry, W. H. Woodall, C. W. Champ and S. E. Rigdon, "A Multivariate Exponentially Weighted Moving Average Control Chart," *Technometrics*, vol. 34, no. 1, pp. 46–53, 1992., doi:10.1080/00401706.1992.10485232
- [30] E. S. Page, "CONTINUOUS INSPECTION SCHEMES," *Biometrika*, vol. 41, no. 1 and 2, pp. 100–115, 1954., doi:10.1093/biomet/41.1-2.100
- [31] R. K. Mehra, "On the identification of variances and adaptive Kalman filtering," *IEEE Transactions on Automatic Control*, vol. 15, no. 2, pp. 175–184, 1970., doi:10.1109/TAC.1970.1099422
- [32] P. J. Huber, "Robust Estimation of a Location Parameter," *Annals of Mathematical Statistics*, vol. 35, no. 1, pp. 73–101, 1964., doi:10.1214/aoms/117703732
- [33] A. P. Sage and G. W. Husa, "Adaptive filtering with unknown prior statistics," in *Proc. Joint Automatic Control Conference*, pp. 760–769, 1969.

BIOGRAPHIES

F. J. Burgos-Flórez received the degree in Mechatronics Engineering in 2013 and the M.Sc. degree in Biomedical Engineering in 2016, both from the Universidad Nacional de Colombia, Sede Bogotá, and the Ph.D. degree in Mechanical Engineering in 2022 from the Universidad del Norte, Colombia. He is currently an Assistant Professor at the Universidad Nacional de Colombia, Sede La Paz, and is recognized as a Junior Researcher by Minciencias. His research focuses on control systems, robotics, and the modeling and simulation of dynamic systems, together with the application of data science and artificial intelligence tools, including the analysis of public and official data sources for engineering and education problems in Colombia. ORCID: <https://orcid.org/0000-0002-5381-1398>

J. G. Popayán-Hernández received the degree in Environmental Engineering in 2010 and the M.Sc. degree in Environmental Engineering in 2015, both from the Universidad Nacional de Colombia, Sede Palmira, and the Ph.D. degree in Environmental Sciences in 2021 from the Universidad del Valle, Colombia. He is with the Departamento de Ciencias Agrícolas, Facultad de Ciencias Agropecuarias, Universidad Nacional de Colombia, Sede Palmira. His research focuses on environmental sciences, environmental technology, and geography, with applications of remote sensing and geographic analysis to land-use change and urban expansion, soil and watershed hydrology, the environmental impacts of agricultural systems, and socio-environmental conflicts in Colombian territories. ORCID: <https://orcid.org/0000-0001-7110-3371>

J. P. Hoyos-Sánchez received the degree in Electronics and Telecommunications Engineering in 2010, the M.Sc. degree in Electronics and Telecommunications in 2016, and the Ph.D. degree in Electronic Sciences in 2018, all from the Universidad del Cauca, Colombia. He is currently an Associate Professor at the Universidad Nacional de Colombia, Sede La Paz. His research focuses on signal processing, artificial intelligence, and convex optimization, with applications of machine learning and data science to the analysis, modeling, and interpretation of information in engineering, education, and other problems of interest in Colombia. ORCID: <https://orcid.org/0000-0002-1844-9127>